Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 8229–8233, August 1995
Biochemistry

The consensus sequence of a major *Alu* subfamily contains a functional retinoic acid response element

GORDON VANSANT AND WANDA F. REYNOLDS*

Sidney Kimmel Cancer Center†, 3099 Science Park Road, San Diego, CA 92121

Communicated by Roy J. Britten, California Institute of Technology, Corona del Mar, CA, May 22, 1995

ABSTRACT ***Alu* repeats are interspersed repetitive DNA elements specific to primates that are present in 500,000 to 1 million copies. We show here that an *Alu* sequence encodes functional binding sites for retinoic acid receptors, which are members of the nuclear receptor family of transcription factors. The consensus sequences for the evolutionarily recent *Alu* subclasses contain three hexamer half sites, related to the consensus AGGTCA, arranged as direct repeats with a spacing of 2 bp, which is consistent with the binding specificities of retinoic acid receptors. An analysis was made of the DNA binding and transactivation potential of these sites from an *Alu* sequence that has been previously implicated in the regulation of the keratin K18 gene. These *Alu* double half sites are shown to bind bacterially synthesized retinoic acid receptors as assayed by electrophoretic mobility shift assays. These sites are further shown to function as a retinoic acid response element in transiently transfected CV-1 cells, increasing transcription of a reporter gene by a factor of ≈35-fold. This transactivation requires cotransfection with vectors expressing retinoic acid receptors, as well as the presence of all-*trans*-retinoic acid, which is consistent with the known function of retinoic acid receptors as ligand-inducible transcription factors. The random insertion of potentially thousands of *Alu* repeats containing retinoic acid response elements throughout the primate genome is likely to have altered the expression of numerous genes, thereby contributing to evolutionary potential.**

The genomes of most higher eukaryotes contain repetitive DNA elements derived from genes transcribed by RNA polymerase (pol) III (1, 2). These interspersed repetitive sequences were initially proposed to represent regulatory networks, allowing the coordinate expression of multiple, unlinked genes (3). Further analysis indicated considerable interspecies variation in these DNA sequences and in their sites of insertions, which argued against a fundamental role in gene regulation and gave rise to the concept that interspersed repetitive sequences are selfish DNA with no function or selective advantage to the organism (4). In support of a regulatory function, recent findings indicate that certain of the primate-specific *Alu* repeats are involved in tissue-specific regulation of nearby genes (5–7).

Alu elements are functional pol III genes with internal A and B box promoter elements and are probably derived from 7SL genes. The *Alu* sequences have been amplified and reinserted throughout the genome by a retroposition process involving a RNA intermediate. A few highly conserved source genes produce the transcripts, which serve as intermediates for retroposon formation (8). During the preceding 30–60 million years of primate evolution, a succession of source genes has given rise to extensive *Alu* subfamilies whose members share a few common diagnostic base changes, indicative of mutations in the parental source gene (8–12). Except for these few base changes, most of the source gene sequence has been conserved throughout this period. An analysis of the known *Alu* sequences indicates that mutations have been strongly suppressed at a number of positions, implying that *Alu* sequences have sequence-dependent functions important for primate evolution (13). One postulated function is that these inserts influence the expression of nearby genes (3). In support of this concept, the data presented here indicate that the consensus sequences of evolutionarily recent *Alu* subfamilies contain binding sites for retinoic acid receptors (RARs), transcriptional regulators that are present in most cell types and play important roles in development and cell growth (14–16).

MATERIALS AND METHODS

The DNA sequence preceding the keratin K18 gene, including the proximal *Alu* element, has been reported (ref. 7 and references therein). Plasmid constructions are detailed in the figure legends. The procedures used for CV-1 cell culture, transfection assays, and gel shift assays have also been described (7, 17–19) and are detailed in the figure legends.

RESULTS AND DISCUSSION

***Alu* Sequences Contain Consensus Hormone Response Elements (HREs).** This study began as an investigation into the role of an upstream *Alu* element in the regulation of the human keratin K18 gene. This *Alu* confers copy number-dependent expression to the K18 gene in transgenic mice, suggesting that it insulates the associated gene from negative effects of sequences surrounding random insertion sites (7). This element is also coincident with a DNase I hypersensitive site, which correlates with K18 transcriptional activity (20). We examined the K18-associated *Alu* element for possible regulatory sites and found several potential binding sites for RARs (Fig. 1*A*). Since the mouse K18 homolog is retinoic acid (RA) inducible in embryonal carcinoma cells (23), as is the human K18 gene (data not shown), this suggested the *Alu* element might be involved in this regulation.

RARs are members of the nuclear receptor superfamily of ligand-activated transcription factors, which also includes receptors for steroid hormones, thyroid hormone, glucocorticoids, and vitamin D (14–16). There are three forms of RARs (RARα, -β, -γ) and three forms of retinoid X receptors (RXRα, -β, -γ). These receptors bind most typically as RAR–RXR heterodimers to two adjacent HREs consisting of variants of the consensus sequence AGGTCA. The consensus HRE sequence shown in Fig. 1*A* was deduced from a compilation of naturally occurring and experimentally derived recognition sequences (14, 21). Many naturally occurring HREs deviate at one or more positions from this motif

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: pol, RNA polymerase; HRE, hormone response element; RA, retinoic acid; DR2, HREs with a spacing of 2 bp; RARE, RA response element; CAT, chloramphenicol acetyltransferase; RAR, retinoic acid receptor; RXR, retinoid X receptors.

*To whom reprint requests should be addressed.
†Formerly San Diego Regional Cancer Center.

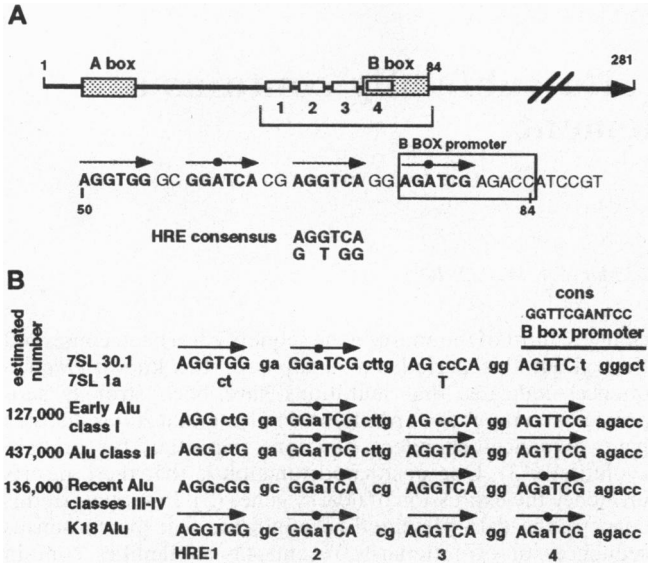

FIG. 1. *Alu* elements contain several consensus HREs. (*A*) A schematic representation of an *Alu* gene (≈281 bp) indicates the relative positions of the A and B box pol III promoter elements and the several potential HREs present in the K18-associated *Alu* sequence. The K18-associated *Alu* sequence including the several HREs (arrows) is shown below. The fourth HRE overlaps with the B box promoter element (boxed). The HREs are arranged as a 2-bp spacing (DR2). The HRE consensus sequence is derived in part from site-selection gel shift assays (21). (*B*) Acquisition of HREs during evolution of *Alu* elements from 7SL genes. The sequences of the HRE region of several *Alu* subfamilies are shown along with the corresponding region of two 7SL genes, 7SL 30.1 and 7SL1a (base differences shown) (22). The *Alu* subfamily consensus sequence for class I (also known as J) is from ref. 11. The consensus sequences for classes II, III, and IV are from ref. 9. Estimated copy numbers for the different subclasses are from ref. 9 and are relative numbers, based on an arbitrary estimate of 750,000 total copies, which is an average of estimates ranging between 500,000 and 1 million. The estimated times of insertions are from ref. 9. The class I consensus is representative of evolutionarily early insertion events (127,000 copies), occurring between 41 and 56 million years ago. The class II consensus represents the majority of inserts with 437,000 copies, inserted 32–57 million years ago. The more recent classes III and IV represent 136,000 copies, inserted 30–40 million years ago. The arrows indicate potential HRE sequences that match the consensus or have one nonconsensus adenine at position three (dot). Residues that fit the HRE consensus are indicated by uppercase letters. Interestingly, the *Alu* B box is a significantly better match for the B box consensus than is the 7SL B box (22).

(14–16). The spacing and orientation of the two half sites is the primary determinant of which nuclear receptor binds to a particular site. RXR also forms heterodimers with thyroid hormone receptor and vitamin D3 receptor, recognizing direct repeats spaced at 4 or 3 bp, respectively (15, 21), whereas RAR–RXR heterodimers recognize direct repeats separated by either 2 or 5 bp (14–16). The several K18-associated *Alu* HREs are arranged as direct repeats with a spacing of 2 bp (DR2) (Fig. 1*A*), which is consistent with binding by RARs.

Comparison of *Alu* consensus sequences with the parental 7SL sequences indicates an evolutionary trend toward the acquisition of multiple *Alu* HREs. The K18-associated *Alu* is of the evolutionarily more recent subfamilies (class III and IV) and has four HRE motifs, designated HRE 1–4 in Fig. 1*B*. The analogous regions of two 7SL genes contain HRE 1 and HRE 2 (7SL30.1) or only HRE 2 (7SL1A) (22), and the HRE 34 pair is a lesser match for the consensus. The *Alu* sequences have been subdivided into classes I–IV (9, 11), reflecting the amount of time elapsed since insertion into the genome, with class I being the oldest *Alu* elements. The consensus sequences (shown in Fig. 1*B*) are made up of those residues that appear most frequently at each position for members of that class and are thus thought to represent the sequence at the time of insertion and therefore the probable sequence of the parental source genes. Class I, the oldest subfamily, is most similar to the 7SL sequence (11) and contains HRE 2 as well as HRE 4, but it lacks two adjacent HREs that fit the consensus with only one base deviation. The class II early *Alu* subfamily represents the majority of *Alu* repeats (estimated relative number of 437,000 copies, based on an estimate of 750,000 total copies; ref. 9). An excellent HRE 3 appears in this subclass, separated from HRE 4 by 2 bp, resulting in one potential DR2 binding site. HRE 2 and HRE 3 remain separated by 4 bp, as in the class I *Alu* and the 7SL sequence. However, in the more recent *Alu* classes, III and IV, a 2-bp deletion between HRE 2 and 3 changed the spacing from 4 to 2 bp, such that there are three direct repeat HREs with a 2-bp spacing (DR2), representing two potential dimer sites, HRE 23 and HRE 34. Certain *Alu* elements within this recent family, such as the K18-associated *Alu*, also contain HRE 1 (present in 7SL30.1 but not class I *Alu* sequences), resulting in four HRE arranged as DR2, representing three potential dimer sites, HRE 12, 23, and 34. Thus, the majority of *Alu* elements (class II, relative copy number ≈ 437,000; ref. 9) contain two adjacent HREs making up one potential DR2 receptor binding site, whereas the more evolutionarily recent *Alu* subfamilies (classes III and IV, relative copy number ≈ 136,000; ref. 9) have at least three HREs, making up two DR2 sites.

The *Alu* HREs Constitute Functional Binding Sites for RARs. To determine if the K18 *Alu* HREs represent functional binding sites for RARs, gel shift assays were performed. Since two adjacent HREs form a binding site for a receptor dimer, double-stranded oligonucleotides were synthesized containing the three possible combinations: HRE 12, HRE 23, and HRE 34 (Fig. 2*A*). The DNAs were end-labeled and incubated with a mixture of bacterially synthesized RARs (RARα and RXRα) (18). Both HRE 23 and HRE 34 bound receptors, producing prominent retarded complexes (Fig. 2*B*, lanes 4 and 6), whereas HRE 12 produced no bound complex (lane 2). A 3-bp substitution in HRE 3 abolished binding to the dimer site HRE 23 (lane 8). Similarly, a 3-bp substitution in HRE 4 essentially abolished binding to HRE 34 (lane 9). We conclude from these results that HRE 23 and 34, but not HRE 12, constitute functional binding sites for these receptors.

In separate experiments, gel shift experiments were performed with either RAR or RXR alone, versus the mixture. For both dimer sites HRE 23 and HRE 34, the mixture produced 5- to 10-fold more bound complex than either receptor alone (data not shown), indicating the heterodimer binds much more effectively to these sites than either homodimer.

The *Alu* HREs Function as a RA Response Element (RARE), Increasing Transcription of a Reporter Gene in Transfected CV-1 Cells. To determine if the *Alu* HREs function as a RARE, reporter constructs were generated consisting of the bacterial chloramphenicol acetyltransferase (CAT) gene fused to the K18 proximal promoter (17), in the absence (XKCAT) or presence (AluXKCAT) of the upstream *Alu* element (Fig. 3*A*). The orientation and position of this *Alu* element is the same as that preceding the native K18 gene. To separate the effect of the HREs from other potential regulatory elements in the *Alu* sequence, a 10-bp mutation was introduced, which abolished the HRE 3 motif, which should eliminate binding to both HRE 23 and HRE 34 (construct MutHRE3). As a second control, to determine the effect of *Alu* gene transcription on CAT gene expression, we tested a construct having mutations in the B box promoter, which were previously shown to abolish transcription by pol III (7) (construct MutBBox). These several reporters were tested by transient transfection in CV-1 cells, in the presence or absence

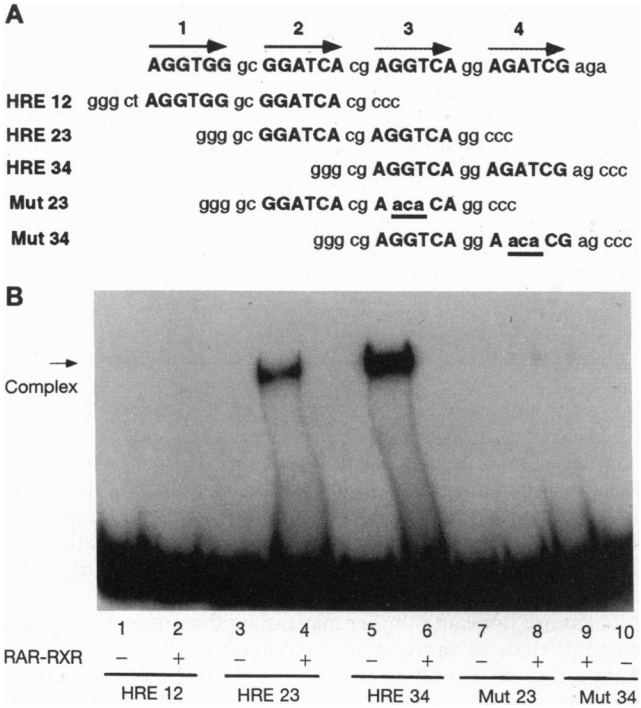

FIG. 2. RARs bind to the *Alu* HREs. (*A*) The sequence of the four potential HRE half sites present in the K18 *Alu* is shown at top. Double-stranded oligonucleotides (top strand shown) containing HRE 12, 23, or 34 and two mutant HREs were synthesized. Mutant HRE 23 contained a 3-bp substitution (lowercase letters, underlined) in HRE 3, whereas mutant HRE 34 had a 3-bp substitution in HRE 4. (*B*) Electrophoretic mobility shift assays show binding by RARs to *Alu* HRE sequences. The double-stranded oligonucleotides were radiolabeled by filling in 3-bp overhangs using the Klenow fragment of DNA polymerase I and [^{32}P]dCTP and purified by elution from 5% polyacrylamide gels. RARα and RXRα were synthesized in bacteria as glutathione *S*-transferase fusion proteins (18) and purified by glutathione-affinity chromatography. A mixture of the two proteins (≈50 ng of each) was incubated with equivalent counts (10,000 cpm, 2–4 ng) of each labeled HRE DR2 element for 30 min at 22°C in a reaction volume of 15 μl containing 10 mM Tris·HCl (pH 7.8), 100 mM KCl, 10% (vol/vol) glycerol, 5 mM $MgCl_2$, 1 mM dithiothreitol, and 2 μg of poly(dI-dC). The protein/DNA mixtures were then electrophoresed in a nondenaturing 5% polyacrylamide gel for 2 hr at 4°C at 200 V in 0.5× TBE (1× is 0.089 M Tris borate, 0.089 M boric acid, and 0.002 M EDTA). The gel was dried and exposed to film. An autoradiograph is shown. The labeled DNAs (indicated at bottom) were electrophoresed in the absence (−) or presence (+) of receptors. The arrow indicates the position of the receptor–DNA complex.

of cotransfected constructs expressing RARα and RXRα (19) (Fig. 3*B*). In the presence of cotransfected receptors and 1 μM all-*trans*-RA, the upstream *Alu* increased CAT expression by ≈35-fold over the level produced by the proximal K18 promoter alone (Fig. 3*B*, set II, lanes 1 and 2). The mutation of HRE 3 abolished most of this enhancer effect, indicating that HRE 23 and/or HRE 34 are required (lane 3). In contrast, mutation of the B box promoter (and HRE 4) resulted in less than a 2-fold decrease in CAT expression (lane 4), demonstrating that transcription of the *Alu* gene is not required for enhancer activity. Transactivation of the CAT gene required cotransfection with vectors expressing RAR and RXR, as well as the presence of all-*trans*-RA; when receptor constructs were cotransfected in the absence of RA, the *Alu* enhanced CAT expression by no more than 2-fold (Fig. 3*B*, set I, lanes 1 and 2). Similarly, in the presence of RA, but the absence of cotransfected receptors, the enhancement was <2-fold (set III, lanes 1 and 2).

The sequence changes that eliminate HRE 3 in mutant construct MutHRE3 are immediately adjacent to the B box

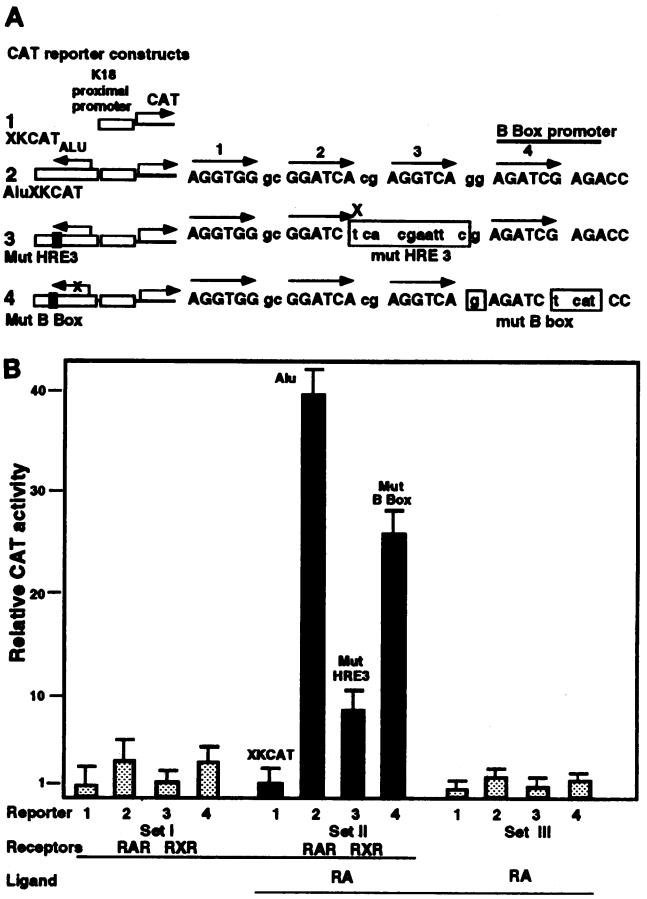

FIG. 3. The *Alu* RARE enhances transcription of a CAT reporter gene in transfected CV-1 cells. (*A*) Schematic representation of reporter constructs. The basal reporter is the previously described XKCATspA, which has upstream sequences from the K18 gene (−251 to +43) fused to the CAT gene (construct 1) (7, 17, 23). Construct 2 (AluXKCAT) has additional K18 upstream sequences (−761 to +43) including the proximal *Alu* gene oriented opposite to the CAT gene. The *Alu* HREs are centered 400 bp upstream of the K18 transcription initiation site. Construct 3 is identical except for a 10-bp mutation, which eliminates HRE 3 and changes the last base pair in HRE 2 (X). Construct 4 contains mutations that render the B box promoter nonfunctional (7) and also changes the final base pair in HRE 4 and the spacing between HRE 3 and 4 from 2 bp to 1 bp. (*B*) Transient transfection assays. CV1 cells were plated at a density of 10^5 cells per 35-mm well (Falcon) 24 hr prior to transfection by a modified calcium phosphate precipitation method according to a protocol described previously (18). The CAT reporter constructs 1–4 are indicated below each lane (4 μg of plasmid DNA per 35-mm^2 well). In sets I and II, 400 ng of pECE-RARα and 100 ng of pECE-RXRα expression vectors (18, 19) were cotransfected along with the CAT reporters. In sets II and III, RA (1 μM all-*trans*-RA) was added to the medium for 24 hr prior to harvest. Cell lysates were prepared and normalized according to protein concentration. Reference plasmids containing the β-galactosidase gene were not used since these have been found to interfere with expression of the *Alu* RARE–CAT reporters. CAT activity was quantitated by a phase-extraction assay (18). The representative data shown are an average of results from three separate experiments.

promoter. It was therefore important to determine if this mutation might inadvertently increase the transcriptional capability of the *Alu* gene, since transcription of a pol III gene has been found to repress nearby pol II gene expression in yeast (24). To compare the transcriptional capability of this HRE 3 mutant to the native *Alu* gene, the constructs were transfected into mouse F9 embryonal carcinoma cells, which lack endogenous *Alu* repeats, and the relative amounts of *Alu* transcription were determined by RNase protection assays (Fig. 4). Mutation of HRE 3 was found to have no effect on *Alu*

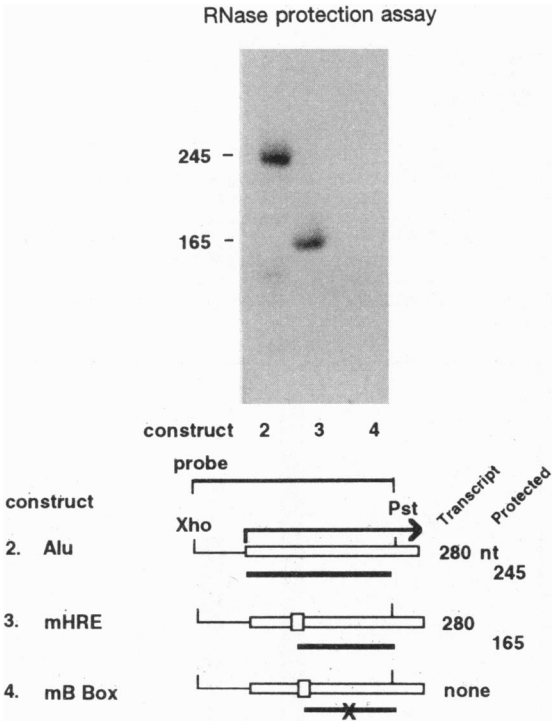

FIG. 4. Mutation of HRE 3 does not affect transcription of the *Alu* gene. Constructs 2–4 (as in Fig. 3*A*) containing the *Alu* gene, the *Alu* gene with the mutated HRE 3, or the *Alu* gene with mutated B box were transfected into mouse F9 embryonal carcinoma cells, which lack endogenous *Alu* sequences. RNA was isolated and hybridized to an *in vitro*-synthesized radiolabeled probe extending from an *Xho* I site 90 bp 5′ of the *Alu* initiation site to a *Pst* I site internal to the *Alu* transcribed region, 245 bp after the initiation site (7). The hybrids were digested with RNase T1, as described (7), and resolved by electrophoresis in 5% polyacrylamide gels containing 8 M urea and 0.5× TBE. An autoradiograph is shown with sizes determined by comparison with markers. The schematic below indicates the predicted sizes of the protected fragments for the three constructs. The native *Alu* transcript protects a 245-nt region of the probe. The transcript of the mutant HRE 3 has a non-base-paired mismatch in the HRE region, which reduces the size of the protected fragment to 165 nt but does not reduce the amount of transcript. The mutated B box promoter abolishes *Alu* transcription, as previously shown (7); no protected fragment was observed.

transcription (lane 3), while mutation of the B box abolished *Alu* transcription (lane 4), consistent with our earlier findings (7). We conclude that the mutation of HRE 3 eliminates the enhancer effect without affecting the transcriptional state of the *Alu* gene.

In summary, these findings indicate that the K18-associated *Alu* contains a functional RARE. Dimer HRE sites (HRE 23 and HRE 34) bind RARs in gel shift assays and function as a RARE in transfected CV-1 cells, increasing expression of a CAT reporter gene by ≈35-fold. Mutation of HRE 3, common to dimer sites HRE 23 and HRE 34, essentially abolished enhancer activity. Mutation of the B box promoter element had relatively little effect, indicating that transcription of the *Alu* by RNA pol III is not essential for enhancer activity. The highest degree of transactivation of the CAT gene required cotransfection with vectors expressing RAR and RXR, as well as the presence of all-*trans*-RA, consistent with the known function of RAR–RXR as a ligand-inducible transcription factor.

The number of *Alu* repeats containing RAREs in the genome is not known. The findings here show that the consensus sequence for evolutionary recent *Alu* classes III and IV, with an estimated copy number of 136,000 (9), contains a RARE. This consensus includes two receptor binding sites, HRE 23 and HRE 34. The HRE 34 motif is also present in the most abundant class of *Alu* elements, class II, with an estimated copy number of ≈437,000. The class II HRE 4 sequence is a better match for the HRE consensus due to one base change (Fig. 1*B*), suggesting that this class of *Alu* elements may also contain a RAR binding site, which would significantly increase the numbers of potential *Alu* RAREs in the genome. Since the consensus is thought to represent the parental source gene sequence, individual *Alu* elements presumably contained RAREs at the time of insertion, but random mutation events occurring since insertion will have eliminated some sites. Based on the 85–89% sequence identity between individual *Alu* elements and the consensus (9), the HRE 23–HRE 34 region would be expected to deviate at approximately three positions in any given *Alu* element. However, since only one of the two dimer sites needs to be retained for receptor binding, it is likely that a significant portion of the class III–IV *Alu* elements retains at least one functional binding site. Moreover, biologically significant *Alu* RARE insertions would not be subject to random mutation rates: If an *Alu* RARE conferred RA inducibility to a pol II gene and the result was advantageous to the organism, that RARE sequence would likely be conserved.

The more relevant number may not be the proportion of *Alu* elements that currently retain functional RAREs but, rather, the number of Alu elements that had RAREs at the time of insertion. Hypothetically, some fraction of *Alu* RAREs will have had immediate effects on expression of nearby pol II genes. *Alu* RARE insertions that resulted in altered gene expression with significant biological consequences will likely have been selected for or against within a few million years, before sufficient time had elapsed to allow random mutations to eliminate the RARE. If so, the sequence of the *Alu* at the time of insertion will determine its primary biological effects. Assuming the consensus sequences represent the source gene sequences, all of the class III–IV *Alu* elements will have had RAREs at the time of insertion.

The random insertion of *Alu* RAREs throughout the primate genome seems hazardous, suggesting mechanisms exist to restrict the function of deleterious RARE insertions. Most significantly, the majority of *Alu* elements are presumed to have inserted into transcriptionally inert, heterochromatic regions of the genome where a RARE would have no effect. Relatively few *Alu* elements would have inserted near enough to a pol II promoter to function as a RARE. Nevertheless, during the preceding 30–60 million years of primate evolution, many *Alu* elements are likely to have inserted near genes for which a proximal RARE was deleterious. Such insertion events would presumably be selected against and thus deleted from the gene pool. Conversely, some fraction of *Alu* elements probably inserted near genes for which RA inducibility was advantageous; the K18 gene is a likely example. Individuals carrying these insertions would be retained in the gene pool. Of the *Alu* RAREs that are currently in the genome, most probably have a neutral effect, and some fraction probably confers a selective advantage.

The coincidence of the B box promoter and HRE 4 suggests another potential regulatory mechanism in which a RAR competes with abundant pol III transcription factors for binding to the HRE region. The B box is bound by the pol III transcription factor TFIIIC, a 500-kDa complex that would effectively block binding by RAR to the several HREs. Interestingly, RA treatment of F9 embryonal carcinoma cells results in a sharp decrease in the amount of pol III transcription factors (25) while inducing the expression of some RARs, conditions favoring RAR binding to available *Alu* RAREs. RAR and the pol III factors could have antagonistic effects, considering the recent finding that active transcription of a pol III gene can inhibit nearby pol II gene expression (24). *Alu* RARE function might also be negatively regulated through

binding by the orphan receptor COUP, which recognizes AGGTCA (HRE 3) and competes for binding by positive-acting RARs (19). Moreover, cell-type-specific factors exist that influence the ability of RAR to bind and transactivate through particular RAREs (26). Finally, *Alu* sequences contain one or more negative regulatory elements, distinct from the HRE region, which can inhibit transcription of a nearby pol II reporter gene (27, 28) and could moderate the RARE effect.

The existence of thousands of RAR binding sites within *Alu* repeats might be expected to deplete soluble receptors and thus interfere with function. However, most *Alu* elements are thought to be sequestered in inaccessible chromatin domains. As evidence for this concept, the overall amount of *Alu* transcription *in vivo* is far below that expected based on the numbers of *Alu* genes, even though individual *Alu* elements can be transcribed *in vitro* using cell-free systems, suggesting the chromatin state of most *Alu* elements *in vivo* blocks transcriptional activity. As further evidence that most *Alu* repeats do not function as free binding sites, the *Alu* sequences contain pol III promoter elements, and yet the large number of *Alu* repeats has no apparent effect on the availability of pol III transcription factors.

Why did *Alu* source genes evolve RAREs? The source genes are required to be transcriptionally active to provide RNA for retroposon formation; thus, the embedded RAREs apparently increase transcriptional capability for the source genes, although there is no evidence that RARs directly activate *Alu* or any other pol III genes. Alternatively, the acquisition of a RARE could indirectly enhance source gene transcription by increasing the probability that a nearby pol II gene will be transcriptionally active. This could promote the assembly of an active chromatin domain, which includes the nearby *Alu* element, making it more accessible to the pol III transcriptional machinery. In support of this concept, the transcriptional activity of particular *Alu* elements has been linked to the transcriptional activity of nearby genes (20, 29).

Other studies have suggested that expression of *Alu* or other interspersed repetitive sequences correlates with tissue-specific expression of associated genes (30) or directly influences the expression of nearby genes (31, 32). Most specifically, an *Alu* element within an intron of the T-cell-specific CD8α gene functions as a T-cell-specific enhancer, having acquired binding sites for several transcription factors present in T cells, including GATA-3 and LyF-1 (5). As a second example, an *Alu* upstream of a gene encoding a T-cell receptor subunit functions as a transcriptional enhancer in T cells (and a repressor in basophils) (6). In both cases, the relevant sequence changes are outside the HRE region and probably appeared after insertion of these particular *Alu* elements. In contrast, the *Alu* RAREs are present in the consensus sequences, indicating their presence in the *Alu* source genes. Since one or more forms of RARs are expressed in most cell types, the acquisition of a RARE would be more likely to benefit a source gene than a T-cell-specific enhancer, since the enhancer would need to function in germ cells to give rise to heritable retroposons.

In conclusion, these findings indicate that the recent *Alu* consensus contains functional RAREs. Probable receptor binding sites also exist within the most abundant class II *Alu* repeats. Considering the large numbers of *Alu* repeats in the primate genome, many genes are likely to have been affected by the insertion of nearby *Alu* RAREs at some time during the preceding 30–60 million years. Accordingly, the random insertion of *Alu* RAREs is likely to have had important consequences during the evolution of primates, generating genomic plasticity by altering the levels of protein expression in response to retinoids.

We thank Magnus Pfahl for RAR and RXR expression constructs, Robert Oshima for plasmids, and Darcy Wilson for a critical reading of the manuscript. This work was supported by a National Institutes of Health Grant RR09118-09 to W.F.R.

1. Weiner, A. M., Deininger, P. L. & Efstratiatis, A. (1986) *Annu. Rev. Biochem.* **55,** 631–661.
2. Howard, B. H. & Sakamoto, K. (1990) *New Biol.* **2,** 759–770.
3. Britten, R. J. & Davidson, E. H. (1969) *Science* **165,** 349–357.
4. Orgel, L. E. & Crick, F. H. C. (1980) *Nature (London)* **284,** 604–607.
5. Hambor, J. E., Mennone, J., Coon, M. E., Hanke, J. H. & Kavathas, P. (1993) *Mol. Cell. Biol.* **13,** 7056–7070.
6. Brini, A. T., Lee, G. M. & Kinet, J.-P. (1993) *J. Biol. Chem.* **268,** 1355–1361.
7. Thorey, I. S., Cecena, G., Reynolds, W. & Oshima, R. G. (1993) *Mol. Cell. Biol.* **13,** 6742–6751.
8. Deininger, P. L., Batzer, M. A., Hutchison, C. A., III, & Edgell, M. H. (1992) *Trends Genet.* **8,** 307–311.
9. Britten, R. J. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 6148–6150.
10. Schmid, C. & Maraia, R. (1992) *Curr. Opin. Genet. Dev.* **2,** 874–882.
11. Jurka, J. & Smith, T. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 4775–4778.
12. Britten, R. J., Bain, W. F., Stout, D. B. & Davidson, E. H. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 4770–4774.
13. Britten, R. J. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 5992–5996.
14. Umesono, K., Murakami, K. K., Thompson, C. C. & Evans, R. M. (1991) *Cell* **65,** 1255–1266.
15. Leid, M., Kastner, P. & Chambon, P. (1992) *Trends Biol. Sci.* **17,** 427–433.
16. Pfahl, M. (1993) *Endocr. Rev.* **14,** 651–658.
17. Oshima, R. G., Abrams, L. & Kulesh, D. (1990) *Genes Dev.* **4,** 835–848.
18. Pfahl, M., Tzukerman, M., Zhang, X.-K., Lehman, J. M., Hermann, T., Wills, K. N., & Graupner, G. (1990) *Methods Enzymol.* **153,** 256–270.
19. Tran, P. Zhang, X.-K., Salbert, G., Hermann, T., Lehman, J. M. & Pfahl, M. (1992) *Mol. Cell. Biol.* **12,** 4666–4676.
20. Neznanov, N. S. & Oshima, R. G. (1993) *Mol. Cell. Biol.* **13,** 1815–1823.
21. Kurokawa, R., Yu, V. C., Naar, A., Kyakumoto, S., Han, Z., Silverman, S., Rosenfeld, M. G. & Glass, C. K. (1993) *Genes Dev.* **7,** 1423–1435.
22. Ullu, E. & Weiner, A. M. (1985) *Nature (London)* **318,** 371–374.
23. Oshima, R. G., Trevor, K., Shevilnsky, L. H., Rynder, O. A. & Cecena, G. (1988) *Genes Dev.* **2,** 505–516.
24. Hull, M. W., Erickson, J., Johnston, M. & Engelke, D. R. (1994) *Mol. Cell. Biol.* **14,** 1266–1277.
25. White, R. J., Stott, D. & Rigby, P. W. J. (1989) *Cell* **59,** 1081–1092.
26. Glass, C. K., Devary, O. V. & Rosenfeld, M. G. (1990) *Cell* **63,** 729–738.
27. Saffer, J. D. & Thurston, S. J. (1989) *Mol. Cell. Biol.* **9,** 355–364.
28. Tomilin, N. V., Igushi-Ariga, S. M. M. & Ariga, H. (1990) *FEBS Lett.* **263,** 69–72.
29. Slagel, V. K. & Deininger, P. L. (1989) *Nucleic Acids Res.* **17,** 8669–8682.
30. Sutcliffe, J. G., Milner, R. J., Gottesfeld, J. M. & Reynolds, W. F. (1984) *Science* **225,** 1308–1315.
31. Carlson, D. P. & Ross, J. (1983) *Cell* **34,** 857–864.
32. Wu, J., Grindlay, J., Bushel, P., Mendelsohn, L. & Allan, M. (1990) *Mol. Cell. Biol.* **10,** 1209–1216.